Optimal Testing for Crowd Workers

Jonathan Bragg  
University of Washington

Mausam  
IIT Delhi

Daniel S. Weld  
University of Washington
Crowdsourcing

- Fun
- Altruism
- Money
Supervised learning

Major Trend #1: Scale and deep learning

- Large neural net
- Medium neural net
- Small neural net
- Traditional algo (SVM, Logistic regression, ...)

Performance vs. Amount of data

Big data
Quality control

High quality

Low quality
Quality control

High quality

Low quality
Quality control in practice

• Gold question insertion
• Worker filtering
Quality control in practice

- Gold question insertion  
  What fraction?
- Worker filtering  
  What accuracy?
Quality control parameters

• Researchers (NLP relation extraction task)
  – 13% gold, 67% accuracy threshold [1]
  – 30% gold, 85% accuracy threshold [2]
  – 20% gold, 80% accuracy threshold [3]

• Industry
  – CrowdFlower: 20% gold, 80% accuracy threshold

Problems

• How to set parameters?
• Static policy is sub-optimal
  – Heterogeneous workers
  – Non-stationary worker population
Goal

Automatically insert gold questions to maximize quality & quantity of individual worker answers
Agent architecture

Give me answers above accuracy $a^*$

Gold questions
Base policy

Agent

Base policy
POMDP policy
Agent architecture

Actions:
- Ask **test** question
- Ask **work** question
- Replace worker

Give me answers above accuracy a*

Gold questions
Base policy

Agent

Base policy
POMDP policy

Environment

Gold (test) questions
Agent architecture

Actions:
- Ask test question
- Ask work question
- Replace worker

Observations:
- Correct test answer
- Incorrect test answer
- Worker leaves

Give me answers above accuracy a*
Gold questions
Base policy

Agent architecture

Environment

Gold questions
Base policy
POMDP policy
Agent architecture

Reward:
1 if correct work answer
-PENALTY if incorrect work answer

Actions:
Ask test question
Ask work question
Replace worker

Observations:
Correct test answer
Incorrect test answer
Worker leaves

Give me answers above accuracy a*
Gold questions
Base policy
Base policy
POMDP policy

Environment

Gold (test) questions
Reinforcement learning

• Challenges
  – Unobservable rewards
  – Exploration should not upset workers

• Our approach
  – Explore with base policy (e.g., “test 20%”)
  – Exploit with learned POMDP policy
POMDP formulation

Worker state (hidden) -> Terminal

Test

Reward: 0

Observation: Correct / incorrect
POMDP formulation

Set PENALTY s.t. Reward > 0 iff worker accuracy > a*

Reward: \( P(\text{correct}) \times 1 + (1 - P(\text{correct})) \times (-\text{PENALTY}) \)

Observation: NULL

Worker state (hidden)

Terminal

Test

Reward: 0

Observation: Correct / incorrect

Work
POMDP formulation

Set PENALTY s.t. Reward > 0 iff worker accuracy > a*

Reward: $P(\text{correct}) \times 1 + (1 - P(\text{correct})) \times (-\text{PENALTY})$

Observation: NULL

Worker state (hidden)

Terminal

Reward: 0

Observation: Correct / incorrect
Worker state transitions

skillful

¬skillful
Worker state transitions

- **diligent**
  - skillful
  - ¬skillful
- ¬**diligent**
Worker state transitions

- Diligent
  - Skillful: Accuracy (skillful)
  - ¬Diligent: Accuracy (¬skillful)

- ¬Diligent
  - ¬Skillful: Accuracy (¬diligent)
POMDP parameters to estimate

• Probability worker leaves
• Probability worker becomes ¬diligent
• Accuracy (skillful)
• Accuracy (¬skillful)
• Accuracy (¬diligent)
• Ratio of skillful to ¬skillful workers
Experiments

• Existing datasets from Mechanical Turk: LinWiki, LinTag, [1] and Rajpal [2]
• Desired accuracy $a^* = 0.85$

Task: Named Entity Linking

“Only two states -- Vermont and **Washington** -- this year joined five others requiring private employers to grant leaves of absence to employees with newborn or adopted infants.”

- Washington, D.C.
- Washington (state)
Policies

• Work-only
• Test-and-replace
  – Test 20%
  – Replace if accuracy < 0.85
• POMDP-RL
  – Base policy: Test-and-replace (above)
  – Exploration budget: 20 workers
LinWiki dataset

Cumulative reward vs. Number of questions asked for different methods:
- **POMDP-RL**
- **Test-and-replace**
- **Work-only**

The graph shows how the cumulative reward changes as the number of questions asked increases for each method.
Rajpal dataset

![Graph showing cumulative reward vs number of questions asked for POMDP-RL, Test-and-replace, and Work-only strategies.](image)
LinTag dataset

Cumulative reward vs Number of questions asked

- POMDP-RL
- Test-and-replace
- Work-only
LinTag worker qualities

Number of questions answered vs. Accuracy

85%
Related work

• Reinforcement learning in educational games
  [Mandel et al. AAMAS ’14, AAAI ’15, ’16]

• Task-centric control for crowdsourcing quality
  [Dai et al. AAAI ’10, ’13; Kamar et al. AAMAS ’12; Lin et al.
   AAAI ’12; Bragg et al. HCOMP ’13]

• Worker-centric control for crowdsourcing engagement
  [Kobren et al. ‘15, Yin & Chen ’15]
Summary

• Problem: automatically insert gold questions to maximize quality & quantity of crowd work
• POMDP-based reinforcement learning algorithm
• Robust to parameter variations (see paper)
• Up to 111% more reward than common policies
• Code available online
Future directions

- Subjective tasks
- Multiple answers per question
- Worker training [1] & engagement

Thanks!

- **Support:** NSF, ONR, WRF/Cable Professorship, Google, Bloomberg
- **Data:** Chris Lin, Shreya Rajpal, Karan Goel, workers on Mechanical Turk
- **Code:**
  
  https://crowdlab.cs.washington.edu/optimal-training-and-testing-for-crowd-workers

![Diagram](attachment:image.png)

Give me answers above accuracy $a^*$